DOCUMENT RESUME

ED 103 488 TM 004 348

AUTHOR Alderman, Donald L.
TITLE Framework for the Educational Evaluation of Computer-Assisted Instruction. Research Memorandum 73-20.
INSTITUTION Educational Testing Service, Princeton, N.J.
REPORT NO ETS-RM-73-20
PUB DATE Oct 73
NOTE 21p.; Part of this paper was contained in a presentation at the Third Annual Frontiers in Education Conference, 1973

EDRS PRICE MF-$0.76 HC Not Available from EDRS. PLUS POSTAGE
DESCRIPTORS Academic Achievement; Affective Objectives; Cognitive Development; *Computer Assisted Instruction; Computer Oriented Programs; Computers; Educational Objectives; Evaluation; *Program Effectiveness; *Program Evaluation
IDENTIFIERS PLATO; TICCIT

ABSTRACT

PLATO and TICCIT, two systems of computer-assisted instruction, are analyzed with a view to determining their educational effectiveness. This report presents an overview of how this will be accomplished, in full awareness that the concept of "effectiveness" has been subject to diverse interpretation. To avoid misunderstandings, the concept is given as much prior elaboration and specification as is now possible. The entire context of this evaluation of educational effectiveness will be clarified by the delineation of four essential issues; the what, who, how and when of the evaluation. Together these basic questions form dimensions of the analytical approach to evaluating the educational effectiveness of PLATO and TICCIT. (Author/DEP)

ED103488

RM-73-20

RESEARCH MEMORANDUM

FRAMEWORK FOR THE EDUCATIONAL EVALUATION OF COMPUTER-ASSISTED INSTRUCTION

Donald L. Alderman

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

PERMISSION TO REPRODUCE THIS MATERIAL BY MICROFICHE ONLY HAS BEEN GRANTED BY
Donald L. Alderman
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER

Part of this paper was contained in a presentation at the Third Annual Frontiers in Education Conference, 1973.

Educational Testing Service
Princeton, New Jersey
October 1973

TM 004 348

BEST COPY AVAILABLE

Framework for the Educational Evaluation of Computer-Assisted Instruction

Donald L. Alderman

Educational Testing Service

Introduction

The PLATO and TICCIT demonstrations will be analyzed with a view to determining their educational effectiveness. The following report presents an overview of how this will be accomplished, in full awareness that the concept of "effectiveness" has been notoriously subject to diverse interpretation. To avoid misunderstandings, we are giving the concept as much prior elaboration and specification as is now possible. As we see it, the entire context of our evaluation of educational effectiveness will be clarified by the delineation of four essential issues, which for the sake of brevity we pose as four questions that must be answered: the what, who, how, and when of the evaluation. Together these basic questions form the dimensions of our analytical approach to evaluating the educational effectiveness of PLATO and TICCIT.

We first identify what areas of inquiry are appropriate to an evaluation of computer-based education. The cognitive and affective domains of student performance and faculty acceptance represent a gross classification of this "what" dimension. For the cognitive domain, our primary concern will be how the students' achievement and behavior have been affected. For the affective domain, we shall seek important information on the effects of an innovative technology through analyzing reaction to computer curricular materials and to the medium itself. The constituents of both domains will be successively refined and ordered in a hierarchical manner. This greater precision and detail will yield indications

of what subtopics and item-stems are relevant to effectiveness. Breadth of coverage will hopefully ensure our objectivity; detail will permit a telling probe of the shibboleths of computer-assisted instruction as reflected by PLATO and TICCIT.

The several audiences concerned with the demonstrations will constitute the "who" dimension of the educational analysis. Because of their differing perspectives and preconceptions, each audience will serve as an important source of data. Depending on the area of inquiry, the sources of critical information will be the educational institution, government, industry, and the public. This probe of different societal sectors will lead not only to the documentation of diverse viewpoints on educational effectiveness, but also to supportive information for the cost and technical analyses. The expression of affective and cognitive impact across vested interests and characteristics is fundamental to the concept of evaluation; this enables us to study attitudes toward computer-assisted instruction and educational practices across a range of viewpoints (e.g., students, instructors, and administrators), and to examine student performance across a range of ability levels and traits. The finer sifting of the dimension will permit us to focus on what comparison groups will be needed and on what groups concerned with the demonstrations must be represented.

The systematic administration of instruments in each area of inquiry and to each appropriate audience requires a plan of how it is to be implemented. How instruments are to be administered is of course determined by available modes of data acquisition. At present these include: tests, student records, questionnaires, online systems, interviews, and observations. Each mode carries particular considerations for instrumentation, ranging from classical test

BEST COPY AVAILABLE

attributes (reliability and validity) to personnel and financial demands. Whether a mode is appropriate for the collection of cognitive and affective data will depend on what detail of analysis is desired for a particular audience.

Precisely when an instrument should be administered is, similarly, an important consideration. Since baseline control depends on measures taken before the introduction of computer-assisted instruction and concurrent control on measures taken during the actual demonstrations, both procedures rely heavily on coordinating measurements with the timing of administrations. Data collection at specified time intervals is also critical to obtaining trend information. We realize the "when" dimension is a particularly sensitive one. We would not want, for example, to force premature judgments or to measure achievement prematurely, nor would we wish to burden respondents with excessive demands upon their time and attention.

Certainly there are interrelationships among the preceding dimensions (i.e., areas of inquiry, audiences, modes of data acquisition, and time of data acquisition) of the educational analysis, as indicated in Figure 1. Although many combinations of components along these dimensions would be inappropriate for our purposes, the individual cells in this initial schematic representation of what, who, and how begin to identify information critical for assessing effectiveness. The accumulation of that information will depend, however, upon our first identifying (1) the area of inquiry; (2) the responding audience; (3) the instrument form used for data acquisition; and (4) the time of administering the instrument. Through such specification, data can be gathered that attest to the educational strengths and weaknesses of the demonstrations.

Insert Figure 1 about here

Dimensions

A. Areas of Inquiry

To capture what impact the demonstrations have had upon cognitive skills and attitudes, we must isolate indicators from the cognitive and affective domains. For the cognitive domain indications of instructional effectiveness will be gathered through various achievement measures, including standardized tests, objective-based tests, and measures of course performance. We shall also supplement information on student achievement with behavioral data on lesson completion, instructional sequence, and other descriptions of cognitive approach toward the subject matter. Activities of students and instructors will serve as indicators of indirect effects beyond instructional outcomes, such as changes in the allocation of effort and time. Further specification on achievement measures will be possible after the elaboration of course objectives.

We shall substantiate and extend information from the cognitive domain through analysis of empirical data collected in the affective domain. The response data relevant to our assessment of educational impact are subsumed under the categories of courseware, role, and appraisal of computer-based education. Courseware refers to the instructional material employed in the demonstrations; the category encompasses content from the subject area, instructional strategy, mode of delivery to the student, and procedures for producing instructional materials. Role refers to the manner in which computer-assisted instruction is used in the school, with special attention to attitudes and receptivity toward that use. Appraisal concerns the priorities and basis for evaluation held by different audiences. These three classifications of inquiry are intended as guidelines in the development of items for the instruments.

BEST COPY AVAILABLE

The delineation of the areas of inquiry from broad concepts to specific indicators is illustrated in Figure 2. As may be seen, ambiguous and sometimes elusive concepts are brought through successive refinements, to the level of item-stems. Priorities in the evaluation are implied by left-to-right positioning along each level of the hierarchy. Since courseware influences cognitive effects and is amenable to review by subject-matter experts and instructional psychologists, a relationship between the cognitive domain and courseware is depicted. Brackets enclose "production" to emphasize that responses in this area pertain only to logistical, not instructional, issues in the development of computer curricular materials. The hierarchical schema presents a partial elaboration of the "what" dimension in the educational analysis. Though incomplete, the portrayed hierarchy is intended to convey both the depth and breadth of inquiry; it also begins to specify priorities for the allocation of our resources. The data analyses will serve to highlight certain information, from our broad coverage, which reveals the strengths and weaknesses of the respective demonstrations.

Insert Figure 2 about here

B. Audiences

Within the context of planning the educational analysis, audiences (see Figure 3) are meant to denote those persons from whom we shall gather responses in the various areas of inquiry. To allow the expression of significant viewpoints we shall include respondents who are representatives of the educational institution, government, public, and industry. Potential data sources in these sectors are the following: (1) students, instructors, and administrators in the cooperating schools and colleges; (2) state and local boards of education

responsible for policies in those schools and colleges; (3) parents of children in the site schools and visitors to the demonstration sites; (4) committees that advise community colleges on the content areas; and (5) manufacturers of system components for PLATO and TICCIT. The focus of our attention during the baseline period is the educational institution, i. ., the cooperating schools and colleges.

Certain characteristics of students, instructors and administrators are relevant to our identifying appropriate audiences and to our establishing proper comparison groups. The first such characteristic is participant status in the demonstrations. For students, participation is determined by enrollment in classes or courses scheduled to use CAI programs; for instructors, by responsibility for the use of CAI in a course or class; for administrators, by implementation of CAI under their auspices. The additional consideration of extent of involvement according to participant criteria depends upon variations within audiences among participants. For those students and instructors who have no direct contacts with CAI another attribute is the similarity between computer-based and conventional instruction. This similarity of courses entails an examination of overlay in instructional objectives and resource materials. Identification of students in, and instructors of courses which parallel computer-assisted instruction permits us to compare instructional outcomes.

Further classifications of respondents within the educational institution relate to the particular district, school, and course or department. This information is required to investigate effectiveness across content areas, schools, and districts, and to recognize natural differences among audiences.

BEST COPY AVAILABLE

Although a breakdown by districts, schools, and courses or departments appears applicable to students and instructors, an administrative position usually satisfies only one of these categories according to the duties of that administrative office.

Insert Figure 3 about here

While the above delineation of the "who" dimension suffices for a preliminary identification, other characteristics are certainly important in establishing comparable or matched groups for analysis. Beyond descriptive information for classification, traits such as student aptitude or instructor experience might influence achievement or attitude. Random assignment, matched groups, and covariance offer means to overcome anticipated control difficulties; which means are employed remains to be arranged although random assignment is certainly the preferred procedure. The available control procedures and the multitude of potential respondents suggest that comparison groups for computer-based and conventional instruction are accessible. In addition, the use of baseline and concurrent comparisons for audience classifications provides a view of difference and change attributable to computer-assisted instruction.

C. Modes of Data Acquisition

The modes for acquiring data for the educational analysis are enumerated as: tests, records, questionnaires, online systems, interviews and observations. Tests in standardized and objective-based form are appropriate for the instrumentation of cognitive inquiries. This information on student achievement and, possibly, aptitude will be complemented by records and online systems. School records might include additional data on achievement and

aptitude by providing course grades and test scores; they may furnish supplemental information on approach by supplying data on school attendance. If we have access to the files and if we find that comprehensive testing programs already exist, the necessity to administer standardized tests will be eliminated. Online systems provide pertinent achievement and behavior information through various evaluations of student progress, and descriptions of student interaction and sequence through courseware. The abundance of online information encompasses: performance on reviews, exercises, and tests; completion of required and optional materials; latencies; time spent at terminal. Systematic observations, especially in the elementary schools, allow the accumulation of behavioral data.

Questionnaires and interviews will constitute the means for data collection in the affective domain. The standard summative or Likert scale is a suitable format for attitude items in questionnaires. A survey instrument in a multiple-choice format is appropriate for gathering information on audience activities, experience, and characteristics. Interviews provide us an opportunity to elaborate attitudes through group and individual discussions. Online systems can possibly be designed to gather attitudinal data from participants and site visitors. The preceding points illustrate the interrelationship of instrumentation with areas of inquiry and audience.

D. Time of Data Acquisition

Another consideration in instrumentation is the time of administration. As mentioned earlier, baseline and concurrent control procedures are to be an integral part of the evaluation. Baseline measures allow us to establish reference points for achievement and attitude before the introduction of

BEST COPY AVAILABLE

computer-assisted instruction to the schools and colleges. When computer curricular materials are in use for student instruction, measures obtained from students and faculty without direct involvement in the demonstrations provide us with relative standards for concurrent comparisons with results from computer-assisted instruction classes. The time selected for instrument administration is also crucial to the use of pre- and posttests as indicators of student achievement.

Of course the area of inquiry, audience, and mode of data acquisition contribute to the specification of timing. Surveys of activities are less sensitive to the "when" dimension than achievement tests are. The time allotted for a respondent to take a standardized test is less flexible than the time allotted to complete an attitude questionnaire. Requests for cooperation must be realistic; we must not intrude upon the audience's other responsibilities.

Interrelationships

While the dimensions of the educational analysis are interrelated, some of the interrelationships are irrelevant to the evaluation. To convey the import of those interrelationships and to elaborate upon the areas of inquiry, we are supplying the following comprehensive outline of potential affective indicators, with references to each dimension. This delineation expands, to the level of item stems, one section of the hierarchical schema for areas of inquiry. The courseware, role, and appraisal of computer-assisted instruction are detailed through successive refinements of their respective components, and through the consideration of other dimensions

A. Courseware

1. Content. The content of instructional materials refers to (1) the scope of the subject matter; (2) the curricular objectives; and (3) the presentation format. The scope of the subject matter is the range and depth of content covered in instruction. Indications of that coverage include: emphasis upon major points; use of examples and illustrations; agreement between content objectives and instructional practice; instructional challenge of exercises and questions. Students' perceptions of the extent of content are influenced by attitudes and anxiety related to the subject area; this situation therefore necessitates the inclusion of items concerning attitude toward the subject matter and anxiety attributed to the difficulty of the material. Since content is relevant to both traditional and computer-based instruction students might respond to questions about the content of traditional instruction at the baseline period. They might respond to questions about the content of traditional and computer-based instruction at the demonstration period. Another source of data about the content of computer-based materials is the participating instructor.

It is doubtful that students are competent to evaluate curricular objectives. Here instructors would be the appropriate respondent. For this subtopic items include: the clarity of courseware objectives; emphasis upon objectives according to a logical framework; and a statement on the curricular specification required by the courseware. The third subtopic for content relates to the presentation of the instructional materials by computer. Questions to be answered are these: Is the format beyond the capabilities of other media? Are there advantages to computer presentation for this content area? Is computer-assisted instruction appropriate for the content?

BEST COPY AVAILABLE

Students and instructors might respond to these presentation inquiries before and after the implementation of CAI to indicate attitudinal impact for a content area. Another important source of data on the content of courseware is the subject-matter expert. Courseware review by experts in the content areas provides judgmental data to complement response data from students and instructors.

2. Instructional Strategy. The sequence and pacing, feedback, management, and component objectives of instructional materials will depend upon the instructional strategy adopted for courseware. Individualization is an evasive concept, but sequence and pacing are amenable to an operational definition which encompasses the following: adaptation to a particular student; imposition of constraints upon learning; resolution of cognitive difficulties; progress through familiar material; encouragement for instructional challenge; detection of curricular or component weaknesses; and use of theoretical and/or practical criteria. With the exception of the last item these points refer to the nature of the interaction between the student and the instructional system. Instructors and participating students represent appropriate audiences for response following courseware implementation. Non-participating students might provide information on sequence and pacing for conventional instruction during the baseline and demonstration periods.

Comparisons between the extent of available feedback in conventional and computer-based settings rely upon student responses. Item-stems relevant to feedback are: provision of clarification and guidance; assistance from comments regarding answers and status; reinforcement effects; maintenance of learner motivation; explanation of objectives; and, beyond the student audience, the incorporation of diagnostic and administrative information for

instructors. As with other attitudinal inquiries, a questionnaire appears to be a suitable form for instrumentation, with the interview reserved for analysis in depth.

The evaluative procedures, instructional man gement, and component objectives inherent to the courseware require review by instructional psychologists. These reviews by experts provide judgmental data to support analytic information on achievement and behavior and on validity and reliability. For evaluative and management procedures judgments are expected to cover the following: clarity and measurability of mastery criteria; agreement with course objectives; validity for supervision, placement, and pacing; conformity with theoretical or practical criteria. For objectives concerned with component repertoires points for review include: consistency with instructional strategy; generality of objectives across content areas; and the theoretical or practical basis for objectives. The instrumentation for expert review necessitates comprehensive courseware documentation for a sample segment, or a brief demonstration with a concomitant explanation of contingencies, accompanied by an appropriate questionnaire or checklist. The time of review is flexible given completed courseware.

3. Delivery. Those attitudinal statements dealing with delivery pertain to the mode of contact, facility of operation, and availability. While facility and availability are incorporated in the technical analysis, these subtopics of delivery are meant to probe the affective reactions of students. Mode of contact entails items for exploring the personal nature of student-computer exchange, the anxiety attributable both to a technological setting and to an authority figure, and student ability to assume motivational responsibility. Facility of operation refers to the psychomotor requisites for system operation,

BEST COPY AVAILABLE

the visual interpretation of displays, complications introduced by other media, and the dramatic effects of the media span upon content. Availability encompasses: distractions from educational purpose introduced by processing lag and by down time, extension of available hours for instruction; accessibility of terminals; and the adequacy of space allotment for terminals. To describe attitudinal changes toward delivery requires an early and late collection of response data during a student's contact with courseware.

4. Production. Of the component areas of courseware, production is most suitable for instrumentation through interviews with instructors and administrators. Among instructors, authors from cooperating schools are an important source of response data on production. The unstructured response format of interviews facilitates the expression of feelings and opinions about the procedures and process for courseware production. These issues concern the appropriate personnel for authoring courseware, means for obtaining professional recognition for developing courseware, and the specifications necessary for computer-based curricula. Other ramifications of production would best be investigated through questionnaires, particularly questions on the availability of instructional materials and programs, and on the standardization of instructional practices. Whether in interviews or questionnaires, items on production are scheduled for administration during the demonstration period. As might be inferred from the above order of consideration, production and delivery are subordinate to content and instructional strategy as priorities in the educational analysis for courseware.

B. Role of CAI

1. Utilization. Is CAI being properly used in the schools? We intend to investigate this question by exploring the extent of, and impact upon,

student and faculty skepticism. If skepticism within the educational institution acts to inhibit the use of CAI, the manifestation and alteration of that resistance become important factors for the evaluation of educational effectiveness. How computer-based education is to be employed covers the manner of use, commitment of instructional time, users' level of education, content areas, and possible corollaries of utilization. These subtopics involve respondents' attitudes on the use of CAI.

Mode refers to the manner of adoption appropriate for this instructional technology. The alternatives range from outright rejection to complete adoption, a range which suggests the advantages of a summative format for attitude assessment. An allied question is whether CAI is perceived as a revolutionary medium or as another instructional resource. Closely related to mode is another subtopic of utilization, time commitment. That is the proportion of student time in a class or course to be spent with CAI. Views on mode and time commitment are to be obtained from students, instructors, and administrators through questionnaires distributed in the baseline and demonstration years. Instructors and administrators might also respond to an item concerning the use of their time required to become familiar with CAI.

Undoubtedly, level of education and content area contribute to determining the proper mode and time commitment for utilization. Through audience classifications according to certain characteristics, information on such interactions is accumulated without repetitious questioning; responses to each subtopic should yield sufficient information. For level of education items concern: the stage of student involvement, elementary school or community college, with CAI; the potential of the computer medium for instructing handicapped, employed, and distantly located students; the point in teacher training for acquaintance

 BEST COPY AVAILABLE

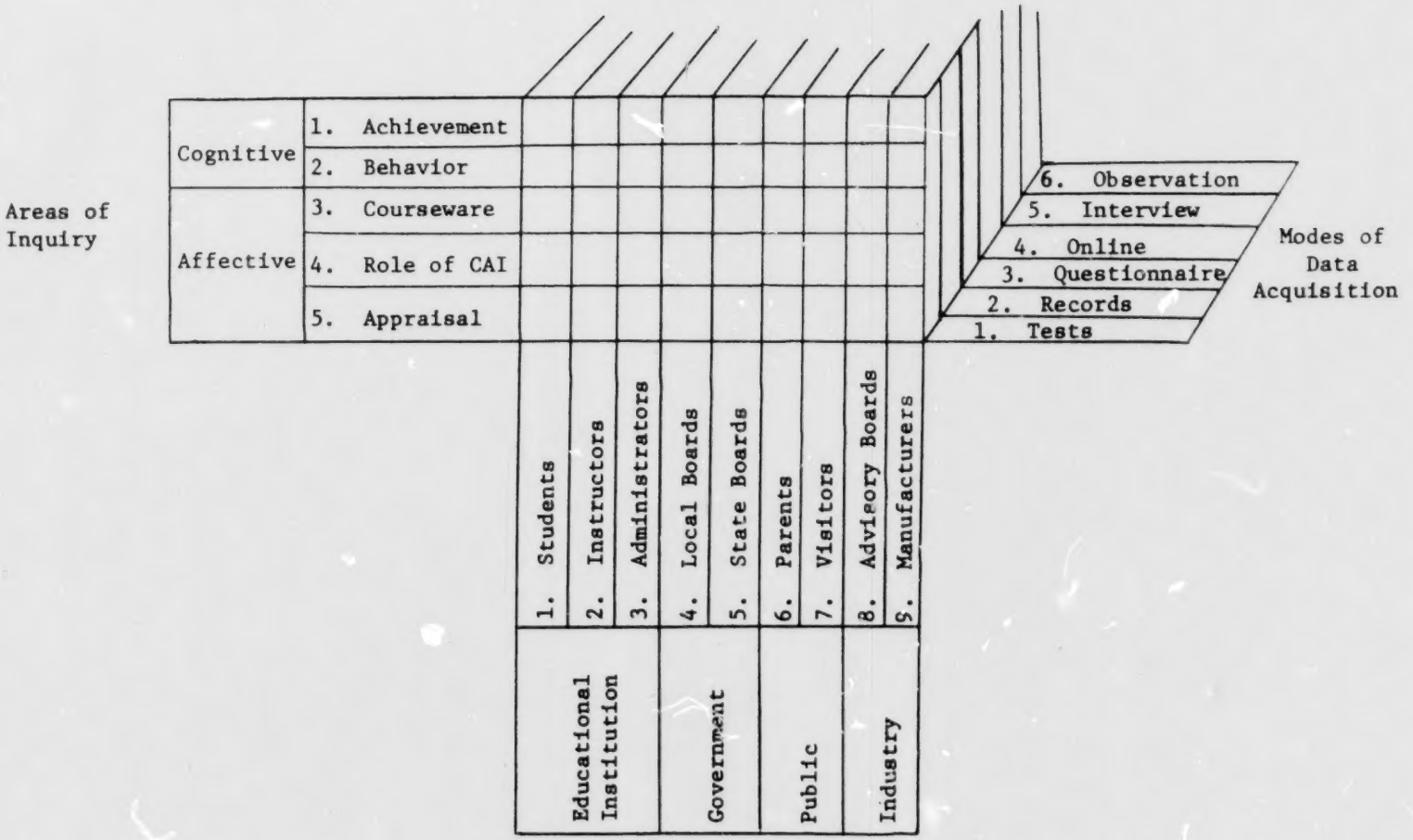
with CAI. For content area items concern application to a subject area and level of that subject suitable for computer-based instruction. Further indicators include possible corollaries of utilization: reduces requisite faculty size; expands the capabilities of instructors; allows more flexibility and variety in student education; makes students more active agents in their education. The audience and instrumentation for level of education and content area follow those for mode and time commitment; corollaries of use might be treated during interviews.

2. Receptivity. Besides providing baseline data as a control reference, receptivity offers an attitudinal reference for utilization through a history of implementation, perception of CAI and conventional education, and concern for innovation. In the attitudinal analysis the meaning of history is limited to the needs, expectations, and responsibilities expressed by participating administrators and instructors. Such an attitudinal history is to be maintained through various contacts with participants, principally through periodic interviews. Items pertinent to perceptions of CAI and conventional education are those that explore the following about computer assisted instruction: it changes role of instructor from purveyor of information to master of resources; it releases faculty time for more profitable pursuits; it neglects the importance of student-faculty contact for modeling; and it detracts from social development by decreasing peer interaction. Although history and perception are oriented toward instructors and administrators, inquiries on concern for innovation include the student audience. Existing attitude scales on concern for innovation are adaptable to our educational analysis. For the role of CAI the indicators for utilization take precedence over those for receptivity in the allocation of resources.

C. Appraisal of CAI

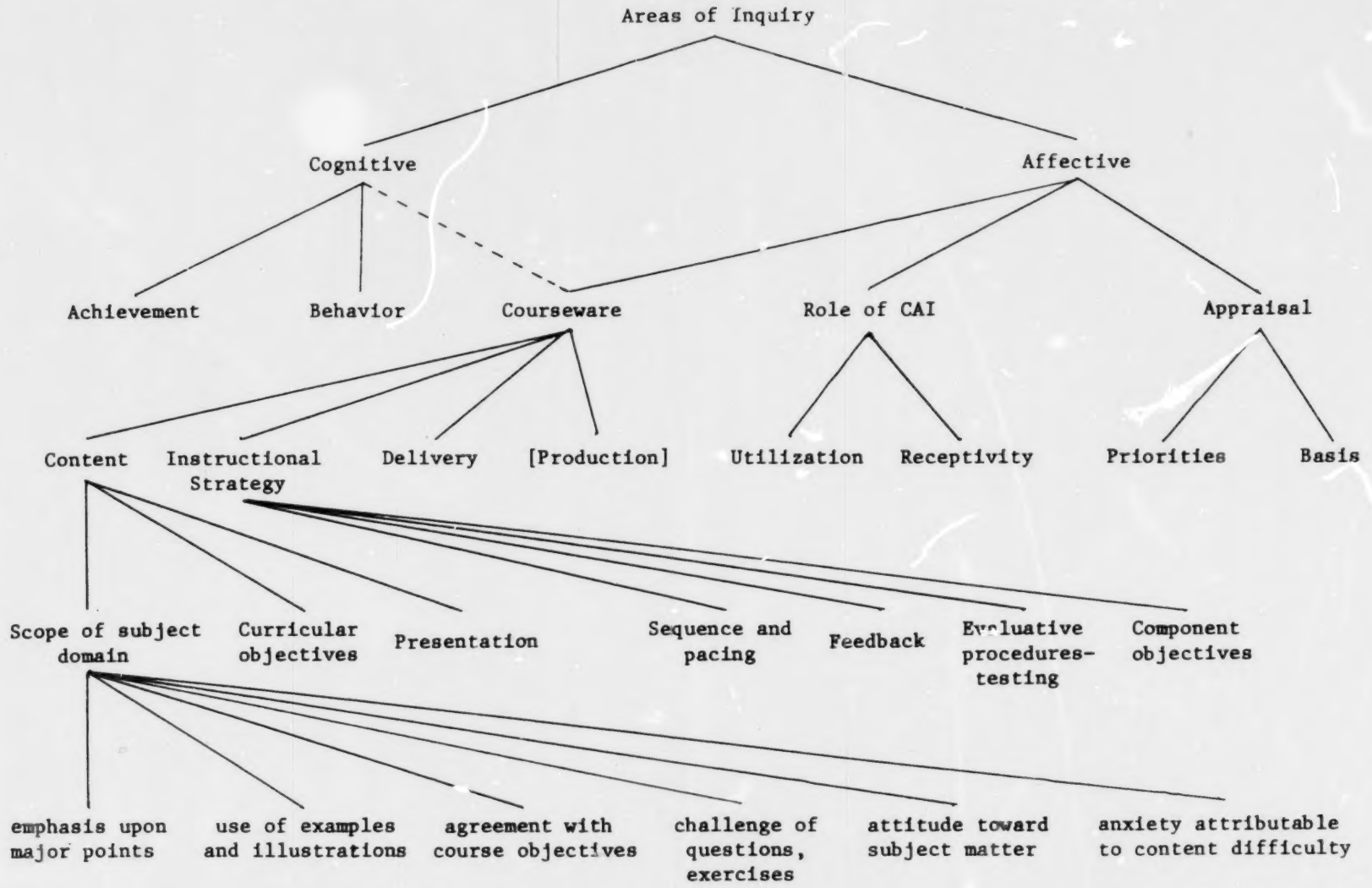
1. Priorities. Since audiences hold different preconceptions and perspectives, it appears reasonable to expect variations in priorities for their appraisal of effectiveness. An audience's priorities should specify the importance which it attributes to cost, technology and education. Both the individual and relative merits of financial, technical, and educational factors in the appraisal of CAI are to be considered by respondents toward the conclusion of the demonstrations. Participants might also indicate their assessments of each factor and their support of CAI. The continuum represented by Likert formats for attitudinal questionnaires fits the purposes of inquiries regarding audience priorities.

2. Basis. Basis for appraisal concerns the previous experience of respondents, and their satisfaction and involvement with conventional education. Statements regarding previous experience are to be accumulated through surveys of activities and behavior, and through school records. Although experience entails matters of fact, subjective responses to technical experience might also indicate comfort with instructional technology. Attitudinal inquiries in this category include indicators of involvement and satisfaction with conventional instruction.



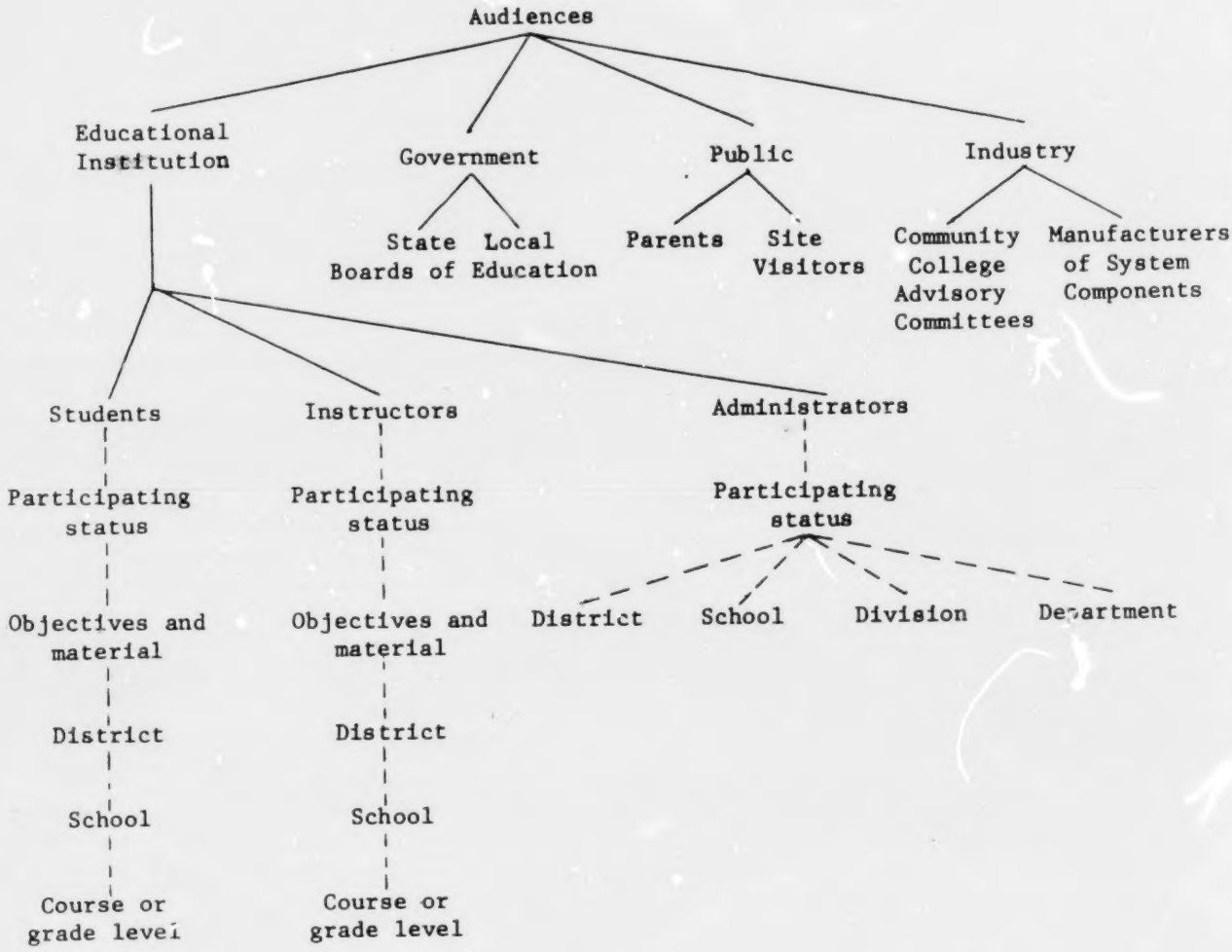
Dimensions of Educational Analysis:
Approach to Effectiveness

Figure 1



Delineation of Areas of Inquiry

Figure 2



Audience Identification

Figure 3

BEST COPY AVAILABLE